

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/523,363
Source: PCT
Date Processed by STIC: 1-26-06

ENTERED



PCT

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/523,363

DATE: 01/26/2006

TIME: 08:51:28

Input Set : A:\Sequence Listing_10523363.txt

Output Set : N:\CRF4\01262006\J523363.raw

3 <110> APPLICANT: National Institute of Advanced Industrial Science and Technology

5 <120> TITLE OF INVENTION: N-glycan synthesis gene

7 <130> FILE REFERENCE: 332-02145

C--> 9 <140> CURRENT APPLICATION NUMBER: US/10/523,363

C--> 10 <141> CURRENT FILING DATE: 2005-02-07

12 <150> PRIOR APPLICATION NUMBER: JP 2002-60679

14 <151> PRIOR FILING DATE: 2002-03-06

16 <160> NUMBER OF SEQ ID NOS: 10

18 <170> SOFTWARE: PatentIn Ver. 2.1

20 <210> SEQ ID NO: 1

21 <211> LENGTH: 1479

22 <212> TYPE: DNA

23 <213> ORGANISM: Homo sapiens

25 <220> FEATURE:

26 <221> NAME/KEY: CDS

27 <222> LOCATION: (1)..(1479)

29 <400> SEQUENCE: 1

```

30 atg gcg gcc ggc gaa agg agc tgg tgc ctg tgc aag ttg ttg agg ttt      48
31 Met Ala Ala Gly Glu Arg Ser Trp Cys Leu Cys Lys Leu Leu Arg Phe
32   1               5               10               15
34 ttt tat tca tta ttc ttc cct ggg ctc att gta tgt gga act tta tgt      96
35 Phe Tyr Ser Leu Phe Phe Pro Gly Leu Ile Val Cys Gly Thr Leu Cys
36           20           25           30
38 gtg tgt ttg gtc att gtc ctt tgg gga atc aga ctg ctg cta cag aga     144
39 Val Cys Leu Val Ile Val Leu Trp Gly Ile Arg Leu Leu Leu Gln Arg
40       35       40       45
42 aag aaa aaa tta gtg tca act agc aaa aat ggg aaa aat caa atg gtg     192
43 Lys Lys Lys Leu Val Ser Thr Ser Lys Asn Gly Lys Asn Gln Met Val
44   50       55       60
46 att gca ttt ttt cat cca tac tgc aat gct ggt gga gga gga gaa aga     240
47 Ile Ala Phe Phe His Pro Tyr Cys Asn Ala Gly Gly Gly Gly Glu Arg
48  65       70       75       80
50 gtt tta tgg tgt gct tta aga gcc ctg cag aaa aag tat cct gaa gca     288
51 Val Leu Trp Cys Ala Leu Arg Ala Leu Gln Lys Lys Tyr Pro Glu Ala
52       85       90       95
54 gtt tat gtt gtt tat acc ggc gat gtt aat gtc aac ggt caa cag ata     336
55 Val Tyr Val Val Tyr Thr Gly Asp Val Asn Val Asn Gly Gln Gln Ile
56   100       105       110
58 cta gaa ggt gct ttc aga aga ttt aac atc aga tta att cac cca gtg     384
59 Leu Glu Gly Ala Phe Arg Arg Phe Asn Ile Arg Leu Ile His Pro Val
60   115       120       125
62 cag ttt gtt ttt tta agg aaa cgc tat ctt gtg gaa gat tca ctg tat     432
63 Gln Phe Val Phe Leu Arg Lys Arg Tyr Leu Val Glu Asp Ser Leu Tyr

```

RAW SEQUENCE LISTING

DATE: 01/26/2006

PATENT APPLICATION: US/10/523,363

TIME: 08:51:28

Input Set : A:\Sequence Listing_10523363.txt

Output Set: N:\CRF4\01262006\J523363.raw

64	130	135	140	
66	cct cac ttc aca ctg ctg ggc caa agt cta gga tcc att ttt ctt ggc	480		
67	Pro His Phe Thr Leu Leu Gly Gln Ser Leu Gly Ser Ile Phe Leu Gly			
68	145	150	155	160
70	tgg gaa gct cta atg cag tgt gtt cct gat gtt tac att gat tca atg	528		
71	Trp Glu Ala Leu Met Gln Cys Val Pro Asp Val Tyr Ile Asp Ser Met			
72	165	170	175	
74	gga tac gct ttt acg ctt cct ctg ttt aag tat ata ggg ggt tgc caa	576		
75	Gly Tyr Ala Phe Thr Leu Pro Leu Phe Lys Tyr Ile Gly Gly Cys Gln			
76	180	185	190	
78	gtt gga agc tat gtt cat tat cct act atc agc acc gac atg ctc tct	624		
79	Val Gly Ser Tyr Val His Tyr Pro Thr Ile Ser Thr Asp Met Leu Ser			
80	195	200	205	
82	gta gtg aag aat caa aat att gga ttt aat aat gca gcc ttc att acc	672		
83	Val Val Lys Asn Gln Asn Ile Gly Phe Asn Asn Ala Ala Phe Ile Thr			
84	210	215	220	
86	agg aat cct ttt ctc agc aaa gta aag ctc atc tac tac tat tta ttt	720		
87	Arg Asn Pro Phe Leu Ser Lys Val Lys Leu Ile Tyr Tyr Tyr Leu Phe			
88	225	230	235	240
90	gct ttt att tat gga ctt gtt ggt tct tgc agt gat gta gtc atg gtc	768		
91	Ala Phe Ile Tyr Gly Leu Val Gly Ser Cys Ser Asp Val Val Met Val			
92	245	250	255	
94	aat tct tct tgg aca cta aac cat att ctc tca cta tgg aaa gtt ggg	816		
95	Asn Ser Ser Trp Thr Leu Asn His Ile Leu Ser Leu Trp Lys Val Gly			
96	260	265	270	
98	aat tgc act aac att gtt tat cca cct tgt gat gtg cag aca ttt ctg	864		
99	Asn Cys Thr Asn Ile Val Tyr Pro Pro Cys Asp Val Gln Thr Phe Leu			
100	275	280	285	
102	gac att ccc tta cat gag aaa aag atg acc cca gga cat ttg ctg gtt	912		
103	Asp Ile Pro Leu His Glu Lys Lys Met Thr Pro Gly His Leu Leu Val			
104	290	295	300	
106	tct gtt ggc cag ttt agg ccg gaa aag aat cat cca ttg cag atc aga	960		
107	Ser Val Gly Gln Phe Arg Pro Glu Lys Asn His Pro Leu Gln Ile Arg			
108	305	310	315	320
110	gcc ttt gct aaa ttg ctg aat aag aag atg gtt gag tca cct cct tcg	1008		
111	Ala Phe Ala Lys Leu Leu Asn Lys Lys Met Val Glu Ser Pro Pro Ser			
112	325	330	335	
114	ctt aaa ctt gtc ctc att gga ggt tgt cgt aac aaa gat gat gaa ctt	1056		
115	Leu Lys Leu Val Leu Ile Gly Gly Cys Arg Asn Lys Asp Asp Glu Leu			
116	340	345	350	
118	agg gta aac caa ctg aga agg ctg tct gag gat tta gga gtt caa gaa	1104		
119	Arg Val Asn Gln Leu Arg Arg Leu Ser Glu Asp Leu Gly Val Gln Glu			
120	355	360	365	
122	tat gtg gaa ttt aaa ata aac att cca ttt gat gaa tta aag aat tat	1152		
123	Tyr Val Glu Phe Lys Ile Asn Ile Pro Phe Asp Glu Leu Lys Asn Tyr			
124	370	375	380	
126	ttg tct gaa gca aca att ggt ctg cat acc atg tgg aac gag cat ttt	1200		
127	Leu Ser Glu Ala Thr Ile Gly Leu His Thr Met Trp Asn Glu His Phe			
128	385	390	395	400

RAW SEQUENCE LISTING

DATE: 01/26/2006

PATENT APPLICATION: US/10/523,363

TIME: 08:51:28

Input Set : A:\Sequence Listing_10523363.txt

Output Set: N:\CRF4\01262006\J523363.raw

```

130 ggg att gga gtt gtg gag tgt atg gca gct ggc aca att atc ctt gca 1248
131 Gly Ile Gly Val Val Glu Cys Met Ala Ala Gly Thr Ile Ile Leu Ala
132          405          410          415
134 cac aat tcg ggg ggc cca aag ctt gac att gtg gtt cct cac gaa gga 1296
135 His Asn Ser Gly Gly Pro Lys Leu Asp Ile Val Val Pro His Glu Gly
136          420          425          430
138 gat ata act ggc ttt ctg gct gag agt gaa gaa gac tat gct gaa act 1344
139 Asp Ile Thr Gly Phe Leu Ala Glu Ser Glu Glu Asp Tyr Ala Glu Thr
140          435          440          445
142 atc gct cac att ctt tcc atg tct gca gaa aag aga ctc caa atc aga 1392
143 Ile Ala His Ile Leu Ser Met Ser Ala Glu Lys Arg Leu Gln Ile Arg
144          450          455          460
146 aaa agt gct cgt gca tct gta agc aga ttc tct gat cag gaa ttt gaa 1440
147 Lys Ser Ala Arg Ala Ser Val Ser Arg Phe Ser Asp Gln Glu Phe Glu
148 465          470          475          480
150 gtg aca ttc cta tca tct gtg gaa aag tta ttt aag taa 1479
151 Val Thr Phe Leu Ser Ser Val Glu Lys Leu Phe Lys
152          485          490
155 <210> SEQ ID NO: 2
156 <211> LENGTH: 492
157 <212> TYPE: PRT
158 <213> ORGANISM: Homo sapiens
160 <400> SEQUENCE: 2
161 Met Ala Ala Gly Glu Arg Ser Trp Cys Leu Cys Lys Leu Leu Arg Phe
162 1 5 10 15
163 Phe Tyr Ser Leu Phe Phe Pro Gly Leu Ile Val Cys Gly Thr Leu Cys
164 20 25 30
165 Val Cys Leu Val Ile Val Leu Trp Gly Ile Arg Leu Leu Leu Gln Arg
166 35 40 45
167 Lys Lys Lys Leu Val Ser Thr Ser Lys Asn Gly Lys Asn Gln Met Val
168 50 55 60
169 Ile Ala Phe Phe His Pro Tyr Cys Asn Ala Gly Gly Gly Gly Glu Arg
170 65 70 75 80
171 Val Leu Trp Cys Ala Leu Arg Ala Leu Gln Lys Lys Tyr Pro Glu Ala
172 85 90 95
173 Val Tyr Val Val Tyr Thr Gly Asp Val Asn Val Asn Gly Gln Gln Ile
174 100 105 110
175 Leu Glu Gly Ala Phe Arg Arg Phe Asn Ile Arg Leu Ile His Pro Val
176 115 120 125
177 Gln Phe Val Phe Leu Arg Lys Arg Tyr Leu Val Glu Asp Ser Leu Tyr
178 130 135 140
179 Pro His Phe Thr Leu Leu Gly Gln Ser Leu Gly Ser Ile Phe Leu Gly
180 145 150 155 160
181 Trp Glu Ala Leu Met Gln Cys Val Pro Asp Val Tyr Ile Asp Ser Met
182 165 170 175
183 Gly Tyr Ala Phe Thr Leu Pro Leu Phe Lys Tyr Ile Gly Gly Cys Gln
184 180 185 190
185 Val Gly Ser Tyr Val His Tyr Pro Thr Ile Ser Thr Asp Met Leu Ser
186 195 200 205

```

RAW SEQUENCE LISTING

DATE: 01/26/2006

PATENT APPLICATION: US/10/523,363

TIME: 08:51:28

Input Set : A:\Sequence Listing_10523363.txt

Output Set: N:\CRF4\01262006\J523363.raw

```

187 Val Val Lys Asn Gln Asn Ile Gly Phe Asn Asn Ala Ala Phe Ile Thr
188      210      215      220
189 Arg Asn Pro Phe Leu Ser Lys Val Lys Leu Ile Tyr Tyr Tyr Leu Phe
190 225      230      235      240
191 Ala Phe Ile Tyr Gly Leu Val Gly Ser Cys Ser Asp Val Val Met Val
192      245      250      255
193 Asn Ser Ser Trp Thr Leu Asn His Ile Leu Ser Leu Trp Lys Val Gly
194      260      265      270
195 Asn Cys Thr Asn Ile Val Tyr Pro Pro Cys Asp Val Gln Thr Phe Leu
196      275      280      285
197 Asp Ile Pro Leu His Glu Lys Lys Met Thr Pro Gly His Leu Leu Val
198      290      295      300
199 Ser Val Gly Gln Phe Arg Pro Glu Lys Asn His Pro Leu Gln Ile Arg
200 305      310      315      320
201 Ala Phe Ala Lys Leu Leu Asn Lys Lys Met Val Glu Ser Pro Pro Ser
202      325      330      335
203 Leu Lys Leu Val Leu Ile Gly Gly Cys Arg Asn Lys Asp Asp Glu Leu
204      340      345      350
205 Arg Val Asn Gln Leu Arg Arg Leu Ser Glu Asp Leu Gly Val Gln Glu
206      355      360      365
207 Tyr Val Glu Phe Lys Ile Asn Ile Pro Phe Asp Glu Leu Lys Asn Tyr
208      370      375      380
209 Leu Ser Glu Ala Thr Ile Gly Leu His Thr Met Trp Asn Glu His Phe
210 385      390      395      400
211 Gly Ile Gly Val Val Glu Cys Met Ala Ala Gly Thr Ile Ile Leu Ala
212      405      410      415
213 His Asn Ser Gly Gly Pro Lys Leu Asp Ile Val Val Pro His Glu Gly
214      420      425      430
215 Asp Ile Thr Gly Phe Leu Ala Glu Ser Glu Glu Asp Tyr Ala Glu Thr
216      435      440      445
217 Ile Ala His Ile Leu Ser Met Ser Ala Glu Lys Arg Leu Gln Ile Arg
218      450      455      460
219 Lys Ser Ala Arg Ala Ser Val Ser Arg Phe Ser Asp Gln Glu Phe Glu
220 465      470      475      480
221 Val Thr Phe Leu Ser Ser Val Glu Lys Leu Phe Lys
222      485      490
226 <210> SEQ ID NO: 3
227 <211> LENGTH: 1581
228 <212> TYPE: DNA
229 <213> ORGANISM: Homo sapiens
231 <220> FEATURE:
232 <221> NAME/KEY: CDS
233 <222> LOCATION: (1)..(1581)
235 <400> SEQUENCE: 3
236 atg gcg gcg ctc aca att gcc acg ggt act ggc aat tgg ttt tcg gct 48
237 Met Ala Ala Leu Thr Ile Ala Thr Gly Thr Gly Asn Trp Phe Ser Ala
238 1 5 10 15
240 ttg gcg ctc ggg gtg act ctt ctc aaa tgc ctt ctc atc ccc aca tac 96
241 Leu Ala Leu Gly Val Thr Leu Leu Lys Cys Leu Leu Ile Pro Thr Tyr

```

RAW SEQUENCE LISTING

DATE: 01/26/2006

PATENT APPLICATION: US/10/523,363

TIME: 08:51:28

Input Set : A:\Sequence Listing_10523363.txt

Output Set: N:\CRF4\01262006\J523363.raw

242		20		25		30											
244	cat	tcc	aca	gat	ttt	gaa	gta	cac	cga	aac	tgg	ctt	gct	atc	act	cac	144
245	His	Ser	Thr	Asp	Phe	Glu	Val	His	Arg	Asn	Trp	Leu	Ala	Ile	Thr	His	
246			35					40					45				
248	agt	ttg	cca	ata	tca	cag	tgg	tat	tat	gag	gca	act	tca	gag	tgg	acg	192
249	Ser	Leu	Pro	Ile	Ser	Gln	Trp	Tyr	Tyr	Glu	Ala	Thr	Ser	Glu	Trp	Thr	
250		50				55				60							
252	ttg	gat	tac	ccc	cct	ttc	ttt	gca	tgg	ttt	gag	tat	atc	ctg	tca	cat	240
253	Leu	Asp	Tyr	Pro	Pro	Phe	Phe	Ala	Trp	Phe	Glu	Tyr	Ile	Leu	Ser	His	
254	65					70				75				80			
256	ggt	gcc	aaa	tat	ttt	gat	caa	gaa	atg	ctg	aat	gtc	cat	aat	ttg	aat	288
257	Val	Ala	Lys	Tyr	Phe	Asp	Gln	Glu	Met	Leu	Asn	Val	His	Asn	Leu	Asn	
258					85					90				95			
260	tac	tcc	agc	tca	agg	acc	tta	ctt	ttc	cag	aga	ttt	tcc	gtc	atc	ttt	336
261	Tyr	Ser	Ser	Ser	Arg	Thr	Leu	Leu	Phe	Gln	Arg	Phe	Ser	Val	Ile	Phe	
262			100					105				110					
264	atg	gat	gta	ctc	ttt	gtg	tat	gct	gtc	cgt	gag	tgc	tgt	aaa	tgc	att	384
265	Met	Asp	Val	Leu	Phe	Val	Tyr	Ala	Val	Arg	Glu	Cys	Cys	Lys	Cys	Ile	
266			115					120				125					
268	gat	gga	aaa	aaa	gtg	ggt	aaa	gaa	ctt	aca	gaa	aag	cca	aaa	ttt	att	432
269	Asp	Gly	Lys	Lys	Val	Gly	Lys	Glu	Leu	Thr	Glu	Lys	Pro	Lys	Phe	Ile	
270		130					135					140					
272	ctg	tcg	gta	tta	ctt	ctg	tgg	aac	ttc	ggg	tta	tta	att	gtg	gac	cat	480
273	Leu	Ser	Val	Leu	Leu	Leu	Trp	Asn	Phe	Gly	Leu	Leu	Ile	Val	Asp	His	
274	145					150				155				160			
276	att	cat	ttt	cag	tac	aat	ggc	ttt	tta	ttt	gga	tta	atg	cta	ctc	tcc	528
277	Ile	His	Phe	Gln	Tyr	Asn	Gly	Phe	Leu	Phe	Gly	Leu	Met	Leu	Leu	Ser	
278				165				170				175					
280	att	gca	cga	tta	ttt	cag	aaa	agg	cat	atg	gaa	gga	gca	ttt	ctc	ttt	576
281	Ile	Ala	Arg	Leu	Phe	Gln	Lys	Arg	His	Met	Glu	Gly	Ala	Phe	Leu	Phe	
282			180					185				190					
284	gct	ggt	ctc	cta	cat	ttc	aag	cat	atc	tac	ctc	tat	gta	gca	cca	gct	624
285	Ala	Val	Leu	Leu	His	Phe	Lys	His	Ile	Tyr	Leu	Tyr	Val	Ala	Pro	Ala	
286		195					200					205					
288	tat	ggt	gta	tat	ctg	ctg	cga	tcc	tac	tgt	ttc	act	gca	aat	aaa	cca	672
289	Tyr	Gly	Val	Tyr	Leu	Leu	Arg	Ser	Tyr	Cys	Phe	Thr	Ala	Asn	Lys	Pro	
290		210					215				220						
292	gat	ggg	tct	att	cga	tgg	aag	agt	ttc	agc	ttt	ggt	cgt	ggt	att	tcc	720
293	Asp	Gly	Ser	Ile	Arg	Trp	Lys	Ser	Phe	Ser	Phe	Val	Arg	Val	Ile	Ser	
294	225					230				235				240			
296	ctg	gga	ctg	ggt	ggt	ttc	tta	ggt	tct	gct	ctt	tca	ttg	ggg	cct	ttc	768
297	Leu	Gly	Leu	Val	Val	Phe	Leu	Val	Ser	Ala	Leu	Ser	Leu	Gly	Pro	Phe	
298			245					250				255					
300	ctg	gcc	ttg	aat	cag	ctg	cct	caa	gtc	ttt	tcc	cga	ctc	ttt	cct	ttc	816
301	Leu	Ala	Leu	Asn	Gln	Leu	Pro	Gln	Val	Phe	Ser	Arg	Leu	Phe	Pro	Phe	
302			260					265				270					
304	aag	agg	ggc	ctc	tgt	cat	gca	tat	tgg	gct	cca	aac	ttc	tgg	gct	ttg	864
305	Lys	Arg	Gly	Leu	Cys	His	Ala	Tyr	Trp	Ala	Pro	Asn	Phe	Trp	Ala	Leu	
306			275					280				285					

VERIFICATION SUMMARY

DATE: 01/26/2006

PATENT APPLICATION: US/10/523,363

TIME: 08:51:29

Input Set : A:\Sequence Listing_10523363.txt

Output Set: N:\CRF4\01262006\J523363.raw

L:9 M:270 C: Current Application Number differs, Replaced Application Number

L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date